Satisfaction Drivers in Retail Banking: Comparison of Partial Least Squares and Covariance Based Methods

Monika Oleksiak*

Submitted: 28.08.2008, Accepted: 19.03.2009

Abstract

The primary goal of the study is to diagnose satisfaction and loyalty drivers in Polish retail banking sector. The problem is approached with Customer Satisfaction Index (CSI) models, which were developed for national satisfaction studies in the United States and European countries. These are multiequation path models with latent variables. The data come from a survey on Poles' usage and attitude towards retail banks, conducted quarterly on a representative sample. The model used in the study is a compromise between author's synthesis of national CSI models and the data constraints.
There are two approaches to the estimation of the CSI models: Partial Least Squares - used in national satisfaction studies and Covariance Based Methods (SEM, Lisrel). A discussion is held on which of those two methods is better and in what circumstances. In this study both methods are used. Comparison of their performance is the secondary goal of the study.

Keywords: satisfaction, loyalty, customer satisfaction index models, banking sector, structural equation models with latent variables, structural equations modeling, partial least squares, covariance based methods

JEL Classification: C13, C39, C51, G21, M31.

*Warsaw School of Economics, e-mail: moleksiak@gmail.com

1 Introduction

In the last decade, Polish retail banks found themselves under growing competitive pressure. For instance, according to the data used in this study, in the first half of 2007, 1.5% of Poles resigned from a bank's services and 3.4% started with a new bank. Customer retention has so become a major issue. The goal of this study is first of all to find out what drives satisfaction and loyalty, which are crucial to keep customers, among Polish retail banks' clients. Customer Satisfaction models constitute a common approach to that problem. First part of the article revises existing Customer Satisfaction models in an attempt to make a useful synthesis. Two methods can serve to estimate these models, both being yet rarely used in Poland. A worldwide discussion is held on which of the two is more appropriate, still without conclusion (Chin (1995), Derquenne, Hallais (2004), O'Loughlin, Coenders (2002), Vilares, Almeida and Coelho (2005)). That is why the secondary goal of the study is to examine what differences in estimates should be expected and to how different recommendations the two methods can lead. The methods are presented in section 2. In the next three sections, the estimated models are verified, then the estimates are compared between methods and, finally, satisfaction and loyalty drivers' hierarchy is presented. The article ends with conclusions and areas for further research.

2 Model development

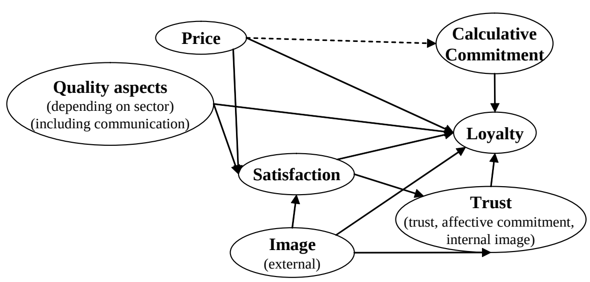
Customer Satisfaction Index (CSI) models were developed for national satisfaction studies in the United States and Western European countries Johnson, Gustafsson, Andreassen, Lervik and Cha (2001). These are multiequation models with latent variables. Latent variables need to be used, because such constructs as satisfaction or loyalty are not directly measurable. Thus, they are measured indirectly, using multiple indicators (symptoms, observed or measured variables), operationalized as 10-point rating scales. These relations are described by the measurement part of the CSI model, while relations between latent variables are held by the structural part of the model. University of Michigan (2005)

2.1 Customer Satisfaction Index (CSI) models

Since the first CSI model, Satisfaction was measured as a cumulative one, not the one concerning the last event. Its typical indicators concern overall satisfaction, fulfillment of expectations and comparison with ideal supplier. Loyalty was first measured purely in behavioral terms - it incorporated declaration of repeated purchase and price sensitivity measurement. This gradually changed to a more affective approach, still asking for repeated purchase, but also for intent to recommend. Johnson, Gustafsson, Andreassen, Lervik and Cha (2001)
The first CSI model was developed in Sweden. In that model, Expectations con-

dition Perceived Value and, together with it - Satisfaction, which in turn generates Complaints, and they both influence Loyalty. The American CSI model added Perceived Quality, conditioned by Expectations and influencing Perceived Value and Satisfaction. Perceived Quality was then split to Perceived Quality of the Product and Perceived Quality of the Service. Models for public sectors were the first to incorporate Trust and Recommendation in Loyalty measurement. The European CSI model added Image as antecedent of Expectations, Satisfaction and Loyalty. The idea was taken from the Norwegian CSI that took into account conclusions made by Johnson, Gustafsson, Andreassen, Lervik and Cha (2001). It denied Complaints - as too rare, Expectations - because the right direction of its influence on Satisfaction is difficult to qualify. Perceived Value - as unclear, was replaced by a pure Price construct. Perceived Quality was replaced by different quality aspects, taken from SERVQUAL methodology. Satisfaction's influence on Loyalty was mediated by Corporate Image, Affective and Calculative Commitment. Johnson, Gustafsson, Andreassen, Lervik and Cha (2001), University of Michigan (2005)
Note that Image is sometimes treated as exogenous, sometimes as endogenous, depending on its understanding. When it results from word-of-mouth and promotional activity ("external image"), it should be exogenous and when it is created by clients experience ("internal image"), it should be endogenous.

Figure 1: A synthesis of structural models used in national CSI studies



The Swiss CSI introduces also Customer Orientation, influencing Perceived Value, Satisfaction, Loyalty and the second new construct - Dialog, that was meant to widen Complaints. Yang, Tian and Zhang (2004) In the European CSI model extension suggested by Ball, Coelho and Machás (2004) this is expanded to Communication, an exogenous variable determining Satisfaction and Loyalty. The idea of Affective Commitment and endogenous understanding of Image is held in Trust, resulting from Communication and Image and influencing Loyalty. Johnson, Gustafsson, Andreassen, Lervik and Cha (2001) recommend also to allow direct influence of all reasonable variables on Loyalty. The author's attempt to make a synthesis resulted in the model presented in Figure 1.

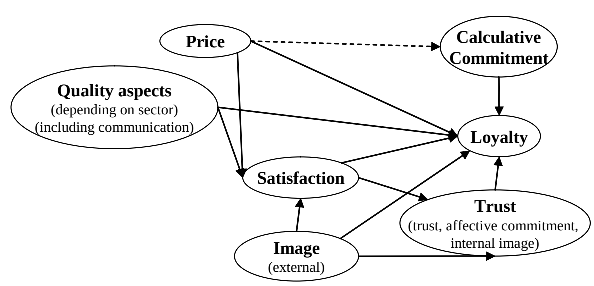
2.2 Data description

The data come from a survey on Poles' usage and attitudes towards retail banks, conducted quarterly by Polish market research company Pentor Research International S.A. on a representative sample. Out of 6092 interviewees surveyed in the first half of 2007, 3119 had an active current account and were qualified to this study. Two parts of the interview are used here. In the first one, respondents are asked to rate their bank's service, employees, account access, interest rates, charges and offer, as well as their satisfaction and intention to continue with that bank. In the second part, image and attitude towards different banks is explored. Respondents are asked to indicate to which banks each of the presented statements is suited. Only the data for the main bank of each customer can be used in the model.

2.3 The model used in the study

The model used in this study is a compromise between the author's synthesis of national CSI models and the data constraints. The Quality aspects available from the data are Offer and Service. Price is renamed to Profitability as banking services do not have one single price, but rather a set of different rates and charges, determining together the overall financial saldo for the customer.

Figure 2: The structural model used in the study



Before any complex modeling could be performed, each latent variable construction was examined by means of the principal components and factor analysis approaches and some indicators were rejected, inter alia all the indicators of Calculative Commitment and thus this latent variable is not in the model. The final measurement model is presented below.

Table 1: The measurement model used in the study

Latent variable name	Indicator variable name	Indicator question
Loyalty	loy1	To what extent would you agree with the statement: "It is unlikely that I would continue to use my main bank services"
	loy2	To what extent would you agree with the statement: "I would willingly recommend my main bank to a friend or a relative"
	loy3	How probable is, that you will use this bank services in the future?
Trust	trust1	Would you point your bank as:... sure and trustworthy
	trust2	...the one, that knows your needs and understands you
	trust3	... the one which services are used by people you respect
	trust4	... the one you like, that is close to you
	trust5	...a bank recommended by friends and relatives
Satisfaction	csi1	To what extent is your main bank similar to your imaginary ideal bank?
	csi2	Taking into account last months, how satisfied are you with your main bank?
	csi3	To what extent would you agree that your main bank is the best for you and there is no better bank in Poland?
	csi4	Taking into account all your expectations and requirements towards a bank, how would you mark your main bank on satisfying those expectations and requirements?
Image	img1	...a sponsor of valuable actions and events
	img2	... having fine commercials
	img3	... a successful bank
	img4	... outstanding
Offer	offer1	How satisfied are you within your bank account with ... range and attractiveness of products and services offered by the bank
	offer2	... the extent to which the offer meets your needs
	offer3	... clarity and comprehensiveness of bank's offer
Service	serv1	... rapidity and efficiency of service in the bank agency
	serv2	... kindness and help provided by bank's employees
	serv3	...access to information about the account balance and history
	serv4	... minimum formalities and requirements needed for opening and using the account
Profitability	profit1	... interests paid for the money you deposited there
	profit2	... tariffs and fees you pay for different services and bank operations

2.4 Hypotheses

Hypotheses regarding the importance hierarchy of Satisfaction and Loyalty drivers are made basing on some psychological common sense with some directions coming from two articles, where estimated models were presented.

In Ball, Coelho and Machás (2004), a model for banking industry based on Portuguese ECSI survey was estimated. According to these results, the ranking of Loyalty drivers, starting with the most important one, is as follows: Satisfaction, Communication, Image, Quality, Complaints, Trust, Expectations and Value. Analogous ranking for Satisfaction drivers is: Image, Quality, Communication, Expectations and Value.

Chatelin, Vinzi and Tenenhaus (2002) estimated an ECSI model for a telecommunication company, which belongs to the third sector of the economy as banking do.

The ranking of Loyalty drivers (Satisfaction, Image, Quality, Expectation, Value and Complaints) is almost the same as in Ball, Coelho and Machás (2004) for the variables present in both models, except for Complaints. The ranking of Satisfaction drivers differs between articles. In Chatelin, Vinzi and Tenenhaus (2002) it is: Quality, Expectation, Image and Value, so only the low position of Value is appearing again.
Assuming that customers value well being, Service should be the most important Satisfaction driver. Some basic rationality assumption should put Profitability on the second place. However customers' Satisfaction is also largely driven by Image, as they tend to identify with the bank's image, so Image may be more important than Profitability. In both articles referenced above, it comes even before Quality and Value is ranked low. Given Poles' weak usage of financial services, Offer should be of least importance. According to the data used in this study, 4.5% of Poles had a credit card, 7.8% had a renewable loan, 10.4% were using standing orders and 10.3% had savings in a bank.
In CSI models Satisfaction is the most important Loyalty driver. Service should be almost as much critical, for bad experience leads to leaving and generates negative word-of-mouth. The place of Trust is difficult to predict, because it should be crucial in that sector, but banks are not really differentiated over that. In Ball, Coelho and Machás (2004) it is the third least important, before Value and Expectations. Image is the second best. However, there are many banks with good image, so it should not prevent churn nor generate intent to recommend, and thus Image is rather awaited as closing the list. Profitability and Offer are expected to be of little importance, because the comparison of banks on that aspects is very difficult. Only the gathering of the information needed for comparison is complicated and time-consuming. And even having all the information, a complex procedure would be necessary to conduct the comparison.

3 The statistical model and estimation procedures

Customer Satisfaction Index models belong to the family of Structural Equation Models with Latent Variables. These models are composed of a measurement model and a structural model.
The measurement model defines relations between latent variables ξ_j, $j = 1, 2, \ldots, J$ (J- the number of latent variables in the model), and their indicators x_{jh}, where j is the index of the latent variable to which the indicator belongs, $h = 1, 2, \ldots, H_j$ (H_j - number of indicators for the j-th latent variable). The measurement model can be reflective or formative. For CSI models, the first notion is appropriate, as it treats latent constructs as existing phenomena "reflected" in the indicators. When the model is formative, the latent variables are just an artificial combinations of the indicators. The reflective measurement model formulation is:

$$x_{jh} = \pi_{jh_0} + \pi_{jh}\xi_j + \varepsilon_{jh}, \tag{1}$$

where π_{jh} - unknown loadings
Usual assumptions for OLS regression and factor analysis on error terms are made, i.e. $E(\varepsilon_{jh}) = 0$, $D^2(\varepsilon_{jh}) = \sigma^2 I$ $(\sigma < \infty)$, $cov(\varepsilon_{jh}, \xi_j) = 0$ and $cov(\varepsilon_{jh}, \varepsilon_{j'h'}) = 0$ for all j, j', h and $h' \neq h$.
The structural model defines the relations between latent variables. It is a usual multiequation model

$$\xi_j = \beta_{j_0} + \sum_{k \in I_j} \beta_{jk} \xi_k + \nu_j, \tag{2}$$

where I_j - collection of the indices of latent variables explaining the j-th latent variable. Usual assumptions for OLS regression and multiequation models on error terms are made, i.e. $E(\nu_j) = 0$, $D^2(\nu_j) = \sigma^2 I$ $(\sigma < \infty)$, $cov(\nu_j, \nu_{j'}) = 0$ for all j and $j' \neq j$.
For both parts of the model, assumptions concerning covariances between error terms or latent variables are not necessary when using covariance based methods, because they can be explicitely modeled.
When indicator variables are standardized, the location parameters π_{jh_0} and β_{j_0} equal 0.
Parameter β_{jk}, called the path coefficient, represents direct effect of the k-th latent variable on the j-th one. Total effect of one latent variable on another is calculated as the sum over all, possibly multistep, paths relating these variables, of products of all direct effects on each path.

3.1 Partial Least Squares

Partial Least Squares (PLS) is the method used in national satisfaction studies (Johnson, Gustafsson, Andreassen, Lervik and Cha (2001)). In the first stage, measurement model parameters and latent variables' individual values are estimated, using an iterative procedure based on the Principal Components approach, but with respect to the structural model dependencies. The second stage is a set of regressions conducted using Ordinal Least Squares, on respective latent variables' estimates. Method description presented below is based on Chatelin, Vinzi and Tenenhaus (2002).
Some data transformations are recommended - in particular, if the scales are not comparable, which is the case here, than indicators are standardized. Latent variables are supposed to be normalized, to assure measurement model identification.
Three steps are iterated:
Step 1. Each standardized latent variable is estimated as a standardized linear combination of its centered indicators. Let us denote by Y_j the external estimate of the standardized j-th latent variable $(\xi_j - \overline{\xi}_j)$($\xi_j$ is normalized by assumption). Y_j is calculated as: $Y_j \propto [\sum_{h=1}^{H_j} w_{jh}(x_{jh} - \overline{x}_{jh})]$, where $\propto$ means standardization. The coefficients w_{jh} after that standardization become $\widetilde{w}_{jh}$, called the outer weights.
Step 2. Each standardized latent variable is estimated as a linear combination of all standardized latent variables being its direct antecedents or consequents, which estimates are taken from the Step 1.

Let us denote by Z_j the internal estimate of the j-th standardized latent variable and let D_j be the set of indicators of those latent variables, which are direct antecedents or consequents of ξ_j. Z_j is calculated as: $Z_j = \sum_{i \in D_j} e_{ij} Y_i$. The coefficients e_{ij} of this combination, called inner weights, are based on the correlation coefficient between the adequate latent variables. They can be calculated using three different schemes: Centroid, Factor Weighting and Path Weighing, giving almost identical results. The latter is most popular, as it tends to lead to the highest squared multiple correlation coefficients (R^2), and it is also used in this study. In the Path Weighing scheme e_{ij} equals the correlation coefficient between Y_j and Y_i, if ξ_i is the consequent of ξ_j. For the antecedents of ξ_j, a multivariate OLS regression is performed of Y_j on those Y_i and e_{ij} is set equal to the respective regression coefficient.
Step 3. Using latent variable estimates from the Step 2, new w_{jh} coefficients are calculated, to be used in the next iteration's Step 1. If the measurement model is reflective, $w_{jh} = \frac{cov(x_{jh}, Z_j)}{var(Z_j)}$, i.e. the new w_{jh} is the regression coefficient in the regression of the indicator variable on the Step 2 estimate of its latent variable.
The iteration procedure starts with w_{jh} set ex. to 1 (depending on the software) and stops when the sum of absolute changes of the outer weights from one iteration to another becomes very small. Final latent variables' estimates are than calculated as:

$$\widehat{\xi}_j = \sum_{h=1}^{H_j} \widetilde{w}_{jh} x_{jh} \tag{3}$$

Having these final latent variables' estimates, structural model parameters can be estimated with OLS. Separate OLS regressions are performed of each endogenous latent variable. Loadings are calculated as: $\widehat{\pi}_{jh} = \frac{cov(x_{jh}, \widehat{\xi}_j)}{var(\widehat{\xi}_j)}$. If observed variables are standardized, this equals to the correlation coefficient between the latent and the observed variable. The algorithm should produce estimates that minimize residual variances in both measurement and structural model. The construction of the Step 2 assures that PLS extracts latent variables in a way to maximize the power of the model to explain endogenous latent variables.
Latent variables' estimates are biased, as they are linear combinations of observed variables that contain measurement error. The bias converges to zero as the number of indicators of the latent variable and the sample size both increase. It is also important that the indicators have little specific variance and much variance common with their latent variable and its other indicators (convergent validity). The common variance (communality) should be at least simply greater than the specific one. Communality is estimated by squared correlation coefficient between the observed and the latent variable, which equals squared loading if indicators are standardized. It should be then grater than 0.5, and consequently loadings should be grater than 0.7. (Ringle, Wende and Will (2005), Rossiter (2002)) In that case the bias is acceptably reduced just at 3 indicators per latent variable. If loadings are only above 0.6, 10 indicators are needed (Chin (1995), Vilares, Almeida and Coelho (2005)). Communalities can

be averaged at latent variable level, giving Average Variance Extracted (AVE). Other measure of construct reliability is Cronbach's alpha and its modification that takes into account different weigh of indicators, called composite reliability. Both should be grater than 0.7. (Rossiter (2002))
It is also needed that each latent variable measures different construct (discriminant validity). It can be verified by checking if each indicator's cross loadings (correlations with latent variables other than its own) are not too high, especially not higher than the loading to its latent variable. Loadings of a latent variable should be also higher than coefficients of paths heading to it. Finally correlations between latent variables should not be too high, especially not higher than latent variable squared AVE. (Ringle, Wende and Will (2005)) If constructs does not hold discriminant validity, collinearity may affect OLS structural model estimates.
Significance of all parameters can be verified by bootstrapping. Special formulas for bootstrapping standard errors are developed to allow statistical comparison of parameters. (Chatelin, Vinzi and Tenenhaus (2002))
Blindfolding is used to cross validate the model. Cross validated communality shows how estimate of the centered latent variable is able to predict its centered symptoms. Cross validated redundancy is similar, but the centered latent variable is replaced by its estimate from the structural model. Both these measures are interpreted in an analogous way as the standard determination coefficients.(Chatelin, Vinzi and Tenenhaus (2002))

3.2 Covariance Based Methods (SEM)

The alternative for PLS are Covariance Based Methods, often referred as SEM or LISREL. The model specification is the same as for PLS, but apart from regression coefficients, latent variables' and error variances and covariances between chosen variables are also model parameters. Modeling explicitly the specific variance of indicators means that SEM relies on Factor Analysis in the measurement model. Variables do not have to be transformed in any way, nor are the latent variables forced to have the variance of 1.
Once the model is stated, with all error terms, regression paths, variances and covariances, free and constrained parameters have to be chosen in order to allow model identification. Naturally, regression path coefficients from errors to the respective variables are constrained to 1. A usual assumption, when estimating structural equation models with latent variables is, for each latent variable, to constrain one of the measurement model coefficients to 1, so setting the measurement scale for this latent variable. With CSI models, this set of assumptions is enough to have an identified model. It is possible to constrain exogenous latent variables' variances to one and by that means have some comparability with PLS, but, as it is still impossible for endogenous latent variables, it is not done here.
The goal of the estimation process in SEM is to find a set of parameters which, for the specified model, imply a theoretical variance-covariance matrix of measured variables,

converging as much as possible to the observed one. Thus, the minimized discrepancy function is:

$$Q(\theta) = (s - \sigma(\theta))^T W^{-1} (s - \sigma(\theta)), \tag{4}$$

where s - vector containing elements of S - the variance-covariance matrix of the observed variables, $\sigma(\theta)$ - vector containing corresponding elements of $\Sigma(\theta)$ - the variance-covariance matrix implied by the model θ - vector of parameters to be estimated, W - weight matrix, which depends on the estimation method.
Three most widely used estimation methods are the Generalized Least Squares (GLS) (with W - the variance-covariance matrix of the residuals) and the Maximum Likelihood (ML) (fitting function: $ln(|\Sigma(\theta)|) - ln(|S|) + tr(\Sigma(\theta)^{-1}) - p$, with p - the number of indicators) and the Asymptotically Distribution Free (ADF) (where W takes into account the skewness of the data).
ML estimators are asymptotically unbiased, consistent and efficient, assuming that the indicators follow a multivariate normal distribution, the sample is large and observations are independent. When multivariate normality assumption is violated, these estimators are no more effective (O'Loughlin and Coenders (2002)). Furthermore, when indicators are skewed, ML estimators present a significant bias (Vilares, Almeida and Coelho (2005)).
A global fit test statistic is based on $Q(\widehat{\theta})$ and is asymptotically chi-squared distributed. Various goodness-of-fit indices have been developed, majority of them based on $Q(\widehat{\theta})$. Values of part of them are normalized to the interval [0,1], with 0.9 as a rule of thumb. Others are not limited and have to be compared with their correspondents for independence and saturated model. Some of them are parsimony adjusted.
Wald statistics are provided to assess the parameters' significance. Lagrange Multiplier-based Modification Indices on the other hand help to find which parameters should be added to the model, but it is recommended to rather start from the most general model and constrain it. Error variances' estimates and squared multiple correlations show which parts of the model should be revised. Any two parameters are compared using the standard z statistic.

3.3 Comparison of estimation methods and hypotheses about their performance

Both these approaches have their limitations and wide discussion is held in the literature on which of them is better for different model specifications and data conditions (Chin (1995), Derquenne and Hallais (2004), O'Loughlin and Coenders (2002), Vilares, Almeida and Coelho (2005)). Here, both are used in imperfect, but very common data conditions, in a way they are usually used by satisfaction researchers. Comparison of their performance is the secondary goal of the study.
PLS produces the more biased estimates the more specific variance the indicators have, and thus many indicators per latent variable are required. The number of indicators per latent variable in this study ranges between 5 and only 2. SEM estimation

with Maximum Likelihood relies on the assumption of the multivariate normal distribution of indicators which, when violated, causes the estimators not to be efficient and all tests to be unreliable. The data used in this study are not normally distributed, or even symmetric, which is usual for rating scales. Both methods' estimates are then expected to be biased, so they will differ. A simulation study by Vilares, Almeida and Coelho (2005) demonstrated that PLS tends to overestimate the measurement model coefficients and underestimate those of the structural one, while SEM shows exactly opposite tendencies. It is thus supposed that the PLS measurement model estimates should be systematically higher than SEM's and that the opposite should be observed for structural model coefficients. As reported by Chin (1995) and Vilares, Almeida and Coelho (2005), PLS tends to flatten the loadings within latent variable, which should also be observed here. Furthermore, as the bias of PLS estimates depends on the measurement error, the difference of measurement model coefficients should be the lower, the less specific variance the indicators have. The matter of interest in this study is to assess the scale of these phenomena.
The author's main hypothesis is that, despite all these differences, the importance hierarchy of Satisfaction and Loyalty drivers is the same, regardless the estimation method.

4 Model quality assessment

Missing and "Don't know" values were imputed with the mean for a given bank, as each method has its own way of missing values treatment and using them would obscure the method-implied differences. Percentage of missing values was below 5%. Then, all observed variables were standardized to keep comparability between both methods. All this was done before the construction of latent variables was examined. Maximum Likelihood estimation was used in SEM, with usual constraints, namely one loading for each latent variable set to 1. This is the only difference between the methods where it is impossible to keep comparability: PLS produces normalized latent variables, while in SEM the endogenous latent variables' variances can not be constrained at all. To cope with that, standardized estimates are analyzed for SEM. An attempt was made to use ADF for SEM, which would be more appropriate with the skewed data, but there was no convergence. Calculations were done in SmartPLS and SPSS Amos.

4.1 PLS model assessment

All constructs exhibit convergent validity. Out of 25 loadings, three are slightly below 0.7, four are between 0.7 and 0.8, fourteen are between 0.8 and 0.9 and four are above 0.9. For most constructs, loadings are also similar. Average Variance Extracted is above 0.5 for all constructs, and even above 0.8 for most of them. The composite reliability is over 0.8 and Cronbach's alpha - over 0.7 for all constructs, except Loyalty.

It is in fact the least, but still, acceptable construct, with loadings of 0.683, 0.698 and 0.808 - two slightly under 0.7 and one outstanding the other two. Consequently, its AVE is only 0.535, composite reliability is 0.774 and Cronbach's alpha - 0.569. The reason is that its indicator questions were in different parts of the questionnaire. The third loading slightly below 0.7 belongs to Service, causing its AVE to be under 0.6.
Discriminant validity conditions are also satisfied for all constructs. All loadings for endogenous variables are much higher than the path coefficient of any path heading to that variable. For all latent variables, squared AVE is higher than the correlation of that variable with any other. And all loadings are higher than respective cross-loadings. However, there are some high cross-loadings: 0.604: img3 → Trust, 0.648: serv4 → Offer, 0.620 and 0.617: offer1 and offer2 (respectively) → Service. It is because these variables' indicator questions were asked in one rotated block. By consequence, there are high correlations between latent variables: 0.435: Profitability with Service, 0.487: Profitability with Offer and the highest: 0.684 Service with Offer. Still, all model parameters are statistically significant, including those that might have been affected by these high correlations. The only parameter with the t statistic slightly below 2 is the one describing the influence of Profitability on Loyalty.
Explanatory power of the model is acceptable, with the R^2 respectively: Satisfaction - 0.404, Trust - 0.431, Loyalty - 0.386. Cross-validated communality is over 0.5, for Satisfaction, Offer, Trust and their indicators. For the key construct i.e. Loyalty, it is rather poor. Cross-validated redundancy for endogenous latent variables amounts for 0.316 - Satisfaction, 0.279 - Trust, 0.192 - Loyalty.

4.2 SEM model assessment

The multivariate normal distribution assumption is not satisfied, because of excess kurtosis. The average specific variance of indicator variables is 0.437, which is quite high. Specific variances over 0.5 are observed for the indicators of Loyalty (0.734, 0.766, 0.564), Service (0.686, 0.656, with other two over 0.4), Image (0.625, 0.594, with other two over 0.3) and Profitability (0.531, the other 0.306). Loyalty and Service raised concerns also in PLS. The remaining specific variances range between 0.164 and 0.449, with the average of 0.339.
All parameters of the model are significant at any reasonable significance level, except the one representing the influence of Profitability on Loyalty, for which the p-value is 0.02. This parameter had also a low t statistic in the PLS model.
Modification Indices point clearly that covariances between Offer, Profitability and Service should not be constrained to zero, which was already seen in PLS. However, in this study they remain constrained, because it is impossible to estimate them in PLS and so freeing them in SEM would lead to incomparability between the methods. This is one of the reasons why the chi-square test rejects the null hypothesis that model-implied and empirical covariance matrices do not differ. Root Mean Square Residual is also half way between independence and saturated model. Still, different goodness of fit indices, not parsimony adjusted, are at the level of 0.85, close to the

rule of thumb of 0.9. Those parsimony adjusted are at the level of 0.77. Root Mean Square Error of Approximation of 0.077 is close to the maximum acceptable value of 0.8, but still below it.
The explanatory power of the model is quite good. The structural model residual variances are 0.369 for Satisfaction, 0.232 for Trust and as little as 0.077 for Loyalty. The R^2 values average to 0.546. For Satisfaction R2 is 0.358 (PLS: 0.404), for Trust it is 0.569 (PLS: 0.431) and for the most important construct - Loyalty, it is 0.634 (PLS: 0.386 only). An interesting finding arises - it seems that the explanatory power of the SEM model is the better compared to the PLS, the worse is the quality of the latent construct. This issue is explored at the end of section 5.1.

5 Final results

Relying on acceptable models, the hypotheses can be verified. Those concerning differences between PLS and SEM estimates are confirmed, with one exception. The importance hierarchy of Satisfaction and Loyalty drivers also agrees with expectations.

5.1 Differences between PLS and SEM estimates

First, the estimates of the measurement model coefficients are indeed all higher in PLS, than in SEM, which is presented in Table 2. The average difference is 0.09. Only in one case PLS and SEM estimates are almost equal. Typically - for 14 out of 25 parameters - the difference is in the (0.03, 0.07) range. For six parameters it lies between 0.1 and 0.15 and for four it is between 0.18 and 0.25. Three highest differences concern Loyalty measurement model parameters, which is related to the bad quality of this construct. Despite the differences, the importance hierarchy of the symptoms for each latent variable construction is the same for both methods, so they produce latent variables having the same meaning. Clearly, the divergence between both methods' estimates is the more important the higher is the specific variance of the indicators. As it can be seen from Table 2, the difference is correlated with the specific variance share, with Pearson correlation coefficient of 0.855 (SEM) and 0.676 (PLS) and Kendall's Tau coefficient of 0.606 (SEM) and 0.450 (PLS). These correlation coefficients are also showing that specific variances are more accurately estimated using SEM. Furthermore, all the indicators, whose parameters estimates' differ by more than 0.1 are those with the SEM specific variance estimate of more than 0.5 and the PLS one of more than 0.4.
In case of Image and Service, the indicators can be split into two groups, as if there were two constructs measured, not one. In fact, the indicator questions bring up two notions. By consequence, one group of indicators has lower communality with the latent variable than the other and the PLS bias is probably larger.
Averaging at the construct level also gives an insight. In Table 3 it can be observed that if there are no indicators with large specific variance, the average difference

Table 2: Measurement model coefficients at indicators level: SEM vs PLS

Parameter	SEM estimate (1)	PLS estimate (2)	Difference (2)-(1) (3)	Specific variance SEM	Specific variance PLS
zloy1 ← Loyalty	0.472	0.698	**0.226**	**0.734**	**0.513**
zloy2 ← Loyalty	0.442	0.683	**0.241**	**0.766**	**0.534**
zloy3 ← Loyalty	0.617	0.808	**0.191**	**0.564**	0.348
zcsi1 ← Satisfaction	0.827	0.895	0.068	0.266	0.199
zcsi2 ← Satisfaction	0.876	0.919	0.043	0.192	0.156
zcsi3 ← Satisfaction	0.779	0.873	0.094	0.336	0.238
zcsi4 ← Satisfaction	0.859	0.912	0.053	0.218	0.169
ztrust1 ← Trust	0.740	0.805	0.065	0.446	0.352
ztrust2 ← Trust	0.782	0.835	0.053	0.382	0.302
ztrust3 ← Trust	0.764	0.818	0.054	0.409	0.331
ztrust4 ← Trust	0.766	0.823	0.057	0.407	0.322
ztrust5 ← Trust	0.792	0.836	0.044	0.366	0.301
zimg1 ← Image	0.637	0.756	**0.119**	**0.594**	**0.429**
zimg2 ← Image	0.612	0.737	**0.125**	**0.625**	**0.457**
zimg3 ← Image	0.820	0.854	0.034	0.327	0.271
zimg4 ← Image	0.820	0.862	0.042	0.328	0.256
zoffer1 ← Offer	0.823	0.889	0.066	0.322	0.209
zoffer2 ← Offer	0.915	0.921	0.006	0.164	0.151
zoffer3 ← Offer	0.759	0.863	0.104	0.424	0.256
zserv1 ← Service	0.758	0.819	0.061	0.426	0.330
zserv2 ← Service	0.743	0.796	0.053	0.449	0.366
zserv3 ← Service	0.561	0.691	**0.130**	**0.686**	**0.523**
zserv4 ← Service	0.586	0.736	**0.150**	**0.656**	**0.458**
zprofit1 ← Profitability	0.685	0.867	**0.182**	**0.531**	0.248
zprofit2 ← Profitability	0.833	0.904	0.071	0.306	0.183

between parameters is the lowest, on the level of 0.06. In the case of Image and Service - constructs with two "subconstructs", it is rather on 0.08-0.09 level. Profitability is not far, also having half of its indicators with large specific variance. Loyalty, which indicators all have low communality, notes the highest average difference. Second, the

Table 3: Measurement model coefficients at indicators level: SEM vs PLS

Construct	Average difference	Average specific variance SEM	Average specific variance PLS	Share of indicators with large specific variance
Loyalty	0.219	0.688	0.465	3/3
Satisfaction	0.068	0.253	0.191	0/4
Trust	0.057	0.402	0.322	0/5
Image	0.093	0.469	0.353	2/4
Offer	0.059	0.303	0.205	0/3
Service	0.081	0.554	0.419	2/4
Profitability	0.126	0.419	0.216	1/2

measurement model coefficients' estimates are really more differentiated in SEM than in PLS. This holds for each latent variable, as can be seen from Table 4. Typically the range of SEM estimates is about twice as wide as the one of PLS, while the variance is two to three times stronger. The most outlying disproportions are observed for Profitability, probably because it has only 2 indicators, in which case PCA generates equal loadings.

Third, the estimates of the structural model coefficients are indeed lower in absolute values in PLS, than in SEM, except for the one of Satisfaction→Trust path. The reason for that exception might be that it is the only path relating inside-model variables, which is favorable in PLS. The difference in this path coefficient estimates is also outstandingly low. Only the one for the path Profitability→Loyalty is lower, but this path coefficient is statistically insignificant. That might suggest that PLS estimates of inside-model paths are much less biased than for the others.

Table 4: Differentiation of measurement model parameters: SEM vs PLS

Construct	Estimates range (Max-Min)		Variance of estimates	
	SEM	PLS	SEM	PLS
Loyalty	0.175	0.125	0.0088	0.0046
Satisfaction	0.097	0.046	0.0018	0.0004
Trust	0.052	0.031	0.0004	0.0002
Image	0.208	0.126	0.0128	0.0042
Offer	0.156	0.059	0.0061	0.0009
Service	0.197	0.128	0.0106	0.0034
Profitability	0.148	0.037	0.0110	0.0007

Table 5: Structural model coefficients: SEM vs PLS

Parameter	SEM estimate (1)	PLS estimate (2)	Difference |(1)|-|(2)| (3)
Satisfaction←Image	0.299	0.240	0.059
Satisfaction←Profitability	0.266	0.200	0.066
Satisfaction←Offer	0.232	0.176	0.056
Satisfaction←Service	0.379	0.176	0.056
Trust←Satisfaction	0.136	0.171	**-0.035**
Trust←Image	0.703	0.578	0.125
Loyalty←Image	-0.212	-0.088	0.124
Loyalty←Profitability	*-0.059*	*-0.031*	*0.028*
Loyalty←Offer	0.184	0.105	0.079
Loyalty←Service	0.295	0.195	0.100
Loyalty←Trust	0.366	0.204	0.162
Loyalty←Satisfaction	0.486	0.376	0.111

For all other structural model parameters, the difference is in the (0.05, 0.17) range. However, it is on average much lower for Satisfaction than for Loyalty heading paths: 0.07 and 0.12 respectively. This results probably from the fact that Satisfaction is a construct of much better quality than Loyalty.
The fact that inside-model coefficients are higher in PLS, while others - in SEM, has a very interesting impact on total effects, which are products of both parameters types. Table 6 presents those total effects that are composed of more than one direct effect. For Offer, Profitability and Service, influencing Trust only through Satisfaction, total difference is reduced almost to zero. For all other relations presented in Table 6, this effect is much weaker as direct effects dominate the relation.

Finally, the relation between SEM and PLS squared multiple correlation coefficients for endogenous latent variables is to be examined. At the end of the section 4.2 it was noted that the SEM ones are the higher, compared to the PLS's, the worse is

Table 6: Total effects composed of more than one direct effect: SEM vs PLS

Parameter	SEM estimate (1)	PLS estimate (2)	Difference (1)-(2) (3)	Deviation $\frac{(3)}{\text{average}((1),(2))}$
Trust←Image	0.743	0.619	0.124	18%
Trust←Profitability	0.036	0.034	**0.002**	**6%**
Trust←Offer	0.032	0.030	**0.002**	**6%**
Trust←Service	0.052	0.048	**0.004**	**8%**
Loyalty←Image	0.206	0.128	0.078	47%
Loyalty←Profitability	0.083	0.051	0.032	48%
Loyalty←Offer	0.308	0.177	0.131	54%
Loyalty←Service	0.498	0.309	0.189	47%
Loyalty←Satisfaction	0.536	0.410	0.126	27%

the quality of the latent construct. In fact, the R^2 ratios order is concordant to the average specific variances one, as presented in Table 7.

Table 7: R-squares and average specific variances: SEM vs PLS

Construct	SEM R^2 (1)	PLS R^2 (2)	R^2 ratio (1)/(2) (3)	Average specific variance	
				SEM	PLS
Satisfaction	0.358	0.404	0.86	0.253	0.191
Trust	0.569	0.431	1.32	0.402	0.322
Loyalty	0.624	0.386	1.64	0.688	0.465

This is another effect of PLS latent variables construction as weighted sums of indicators incorporating measurement error. PLS thus explains the construct and its error with other constructs and their errors, while SEM explains the "pure" construct with other "pure" constructs. It is thought expectable that the lower the measurement errors, the better explanatory power should PLS exhibit compared to SEM.

5.2 Satisfaction and Loyalty drivers

It was confirmed that the importance hierarchy of Satisfaction and Loyalty determinants is the same regardless the estimation method. PLS shows, however, that some parameters' estimates are not statistically different. In SEM all interesting differences are statistically significant, however SEM tests differences of the unstandardized estimates.

Satisfaction drivers hierarchy is presented in Table 8.

The most important Satisfaction driver is Service, as it was expected. Image holds in fact a strong second position before Profitability, proving that being linked with an attractive brand is more important than real money (or at least that it is easier to feel and verify than the financial attractiveness of a bank offer). Offer itself is the least important, as it was anticipated. In SEM all these parameters are statistically different from each other, whilst PLS reports statistically significant differences only

Table 8: R-squares and average specific variances: SEM vs PLS

Construct	Direct Effect	
	SEM	PLS
Service	0.379	0.278
Image	0.299	0.240
Profitability	0.266	0.200
Offer	0.232	0.176

between Service and Offer, Service and Profitability and between Image and Offer.

Table 9: Loyalty drivers - direct and total effects: SEM vs PLS

Construct	Direct Effect		Total Effect	
	SEM	PLS	SEM	PLS
Satisfaction	0.486	0.376	0.536	0.410
Service	0.295	0.195	0.498	0.309
Trust	0.366	0.204	0.366	0.204
Offer	0.184	0.105	0.308	0.177
Image	-0.212	-0.088	0.206	0.128
Profitability	-0.059	-0.031	0.083	0.051

Out of Loyalty drivers Satisfaction reveals to be the most important in both models, as in all CSI models. Trust has the second strongest direct effect, but PLS reports it does not differ significantly from Service's one. In terms of the total effect, Service is even more important, so it is the second most important driver, as it was expected. Trust is situated right in the middle. The third important is Offer (both directly and in total), which is surprising, given its low importance in determining Satisfaction. Direct effect of Profitability is statistically insignificant in both models. As it was already noted, Image direct effect is significantly negative, due to the fact that its positive influence goes through Trust and Satisfaction. When analyzing total effects, both Image and Profitability have positive significant role, Image being more important. As it was anticipated, it almost closes the list. The case of Image's impact on Loyalty shows how important it is to base findings on total effects, not on direct ones.

6 Conclusions and further research

Estimated models provide concordant recommendations concerning Satisfaction and Loyalty drivers: Service is the area that should be regarded with special care, as it is the main driver for both Satisfaction and Loyalty. Offer and Profitability can be equally treated, even with less respect for the latter. That is good information for banks, as it shows that they do not have to compete on pricing. Perhaps less money can be invested in Image creation, for it is very important for Satisfaction, but does not affect Loyalty as much. Clearly Service is a better area for investment, however Image can be helpful in building Trust, which is important for Loyalty.

The results concerning the differences between PLS and SEM estimates are concordant with those obtained by Vilares, Almeida and Coelho (2005) and the postulated hypotheses are confirmed. The conclusion for model builders, asking which method is better, is that if the measurement model is well specified and the questionnaire well designed, then indicators should have little specific variance and by consequence SEM and PLS estimates would differ by less than 0.1 being both very close to the real parameters' values. And even if the measurement model is not perfect, the overall picture and the recommendations are the same.
The model presented can be improved. Both methods of estimation signalize high correlation between Offer, Profitability and Service. This can be resolved by explicitly modeling covariances between these constructs, which is impossible in PLS and thus is not done in this study, but constitutes a field for further research. Second area for development is to assess other Covariance Based Methods, especially Asymptotically Distribution Free Estimation as it does not assume any specific distribution of the data. A model with the covariances explicitly modeled and estimated by ADF would probably be much better than the one described here.
The conceptual model can be developed. It should still base on the synthesis model developed here, but, having dedicated data, it could include more drivers, Calculative Commitment construct, and better Loyalty measurement.
Three phenomena were remarked here that need further investigation. First, if the specific variance of indicators is crucial, the way of measuring it is an issue. It was observed here that the estimates of specific variance from SEM better pointed the indicators with largest differences between PLS and SEM loadings, than the estimates from PLS. If that reveals to be regular, it would be recommended for researchers using PLS to estimate their model in SEM in order to assess the amount of specific variance. Second hypothesis to be verified is whether it is regular that inside-model paths are estimated higher in PLS than in SEM, while other paths are not. Are they less biased then? Finally, it was noted that the SEM squared multiple correlations for endogenous latent variables are the higher compared to the PLS's, the worse is the quality of the latent construct. It seems to result from the methodology, but requires a more in-depth analysis.

References

[1] Ball D., Coelho P. S., Machás A., (2004), The Role of Communication and Trust in Explaining Customer Loyalty: An Extension to the ECSI Model, *European Journal of Marketing* 38

[2] Chatelin Y. M., Vinzi V. E., Tenenhaus M., (2002), State-of-Art on PLS Path Modeling. Through the Available Software, *Les Cahiers de Recherche*, n 764, Groupe HEC

[3] Chin W. W., (1995), Partial Least Squares is to LISREL as Principal Components Analysis is to Common Factor Analysis, *Technology Studies 2*, University of Calgary

[4] Coelho P. S., Esteves S. P., (2007), The Choice Between a 5-point and a 10-point Scale in the Framework of Customer Satisfaction Measurement, *International Journal of Market Research* 49

[5] Derquenne Ch., Hallais C., (2004), Une Méthode Alternative a l'approche PLS : Comparaison et Application aux Modeles Conceptuels Marketing, *Revue de Statistique Appliquée* 52

[6] EPSI Rating Editorial Board (2005), *Pan European CSI Report 2005*

[7] Hoyle, R.H. (ed.) (1995). *Structural Equation Modeling*, SAGE Publications, Inc. Thousand Oaks, Canada.

[8] Johnson M. D., Gustafsson A., Andreassen T. W., Lervik L., Cha J., (2001), The Evolution and Future of National Customer Satisfaction Index Models, *Journal of Economic Psychology* 22

[9] O'Loughlin Ch., Coenders G., (2002), Application of the European Customer Satisfaction Index to Postal Services. Structural Equation Models versus Partial Least Squares, *Working Papers of the Department of Economics*, University of Girona, Girona

[10] Ringle Ch. M., Wende S., Will A., (2005), SmartPLS, University of Hamburg, Hamburg

[11] Rossiter J. R., (2002), The C-OAR-SE Procedure for Scale Development in Marketing, *International Journal of Research in Marketing* 19

[12] Tenenhaus M., (1999), L'approche PLS, *Revue de statistique appliquée* 47

[13] University of Michigan (2005), *American Customer Satisfaction Index Methodology Report*, National Quality Research Center, Stephen M. Ross School of Business, Michigan

[14] Vilares M. J., Almeida M. H., Coelho P. S., (2005), Comparison of likelihood and PLS estimators for structural equation modeling. A simulation with customer satisfaction data, *4th International Symposium on PLS and Related Methods*, Barcelona, Spain, 7-9.09

[15] Wold H., (1985), Partial Least Squares, [in:] *Encyclopedia of Statistical Sciences*, vol.6, [eds.:] Kotz S., Johnson N.L., John Wiley & Sons, New York, 581-591

[16] Yang X., Tian P., Zhang Z., (2004), A Comparative Study on Several National Customer Satisfaction Indices (CSI), *4th Asia Academy of Management Conference, Shanghai*, 16-18.12